3-1-2004

VSP Traveltime Inversion: Near-Surface Issues

Geoff J.M. Moret
Pennsylvania State University

William P. Clement
Boise State University

Michael D. Knoll
Boise State University

Warren Barrash
Boise State University

This document was originally published by Society of Exploration Geophysicists in Geophysics. Copyright restrictions may apply. DOI: 10.1190/1.1707053
Case History

VSP traveltime inversion: Near-surface issues

Geoff J.M. Moret*, William P. Clement†, Michael D. Knoll‡, and Warren Barrash‡

ABSTRACT

P-wave velocity information obtained from vertical seismic profiles (VSPs) can be useful in imaging subsurface structure, either by directly detecting changes in the subsurface or as an aid to the interpretation of seismic reflection data. In the shallow subsurface, P-wave velocity can change by nearly an order of magnitude over a short distance, so curved rays are needed to accurately model VSP traveltimes. We used a curved-ray inversion to estimate the velocity profile and the discrepancy principle to estimate the data noise level and to choose the optimum regularization parameter. The curved-ray routine performed better than a straight-ray inversion for synthetic models containing high-velocity contrasts. The application of the inversion to field data produced a velocity model that agreed well with prior information. These results show that curved-ray inversion should be used to obtain velocity information from VSPs in the shallow subsurface.

INTRODUCTION

A vertical seismic profile (VSP) records energy traveling from a surface source to receivers in a borehole. We can invert the arrival times of compressional or P-waves at the receivers to estimate the subsurface velocity profile near the borehole. In shallow geophysics, these velocity profiles are used to assign depths to reflectors identified on seismic sections (e.g., Allison and Schieck, 1996; Jarvis and Knight, 2000) or to map changes in subsurface properties (e.g., lithology, porosity, pore fluid) that are of interest in environmental or engineering investigations (e.g., Michaels and Barrash, 1997; Milligan et al., 2000).

Many VSP inversions (e.g., Stewart, 1984; Schuster, 1988; Lizarralde and Swift, 1999) assume that the raypaths between sources and receivers can be approximated by straight lines. Schuster et al. (1988) found that the straight-ray assumption is reasonable if velocity contrasts are small and the horizontal offset between the shot and the receivers is less than the depth of the first receiver. In VSPs collected in shallow investigations, however, source offsets from the well axis are commonly of the same order as the depths of the first few receivers. Also, velocities in the shallow subsurface typically span a large range (e.g., Miller and Xia, 1998): from less than 300 m/s to greater than 2000 m/s for P-waves. Large velocity contrasts result in significant refraction of rays at interfaces. For example, since saturated sediments have a much higher P-wave velocity than unsaturated sediments, P-wave raypaths refract sharply at the water table. As a result, a curved-ray model will better approximate the true P-wave raypaths and give more realistic traveltimes than can a straight-ray model in shallow subsurface environments.

One way to estimate velocity from traveltimes using curved rays is layer stripping (Milligan et al., 2000). In this method, the velocity in the first layer is estimated assuming a straight raypath from the source to the first receiver. The interval velocities in subsequent layers are calculated based on the time required for a curved ray to reach the receiver in that layer. Layer stripping can produce good results, but the velocity model is highly sensitive to errors in traveltime picks, as the errors are propagated into the subsequent layer velocities.

We have designed and implemented a curved-ray traveltime inversion for shallow VSP data. Previous curved-ray VSP inversions (e.g., Pujol et al., 1985; Lee, 1990; Mao and Stuart, 1997) were designed to find the depths and velocities of a few discrete layers. As a result, these inversions incorporated reflection data and considered refraction at a small number of layer...
boundaries. Previous curved-ray inversions also lacked quantitative error analysis. Shallow geophysics is often concerned with small fluctuations in subsurface properties, so our inversion uses a large number of layers to image gradual changes in seismic velocity. We estimate the traveltime error from the data using Morozov’s discrepancy principle, and we use this error to estimate the uncertainty in the layer velocities. After describing our inversion, we will demonstrate its application with both synthetic and field experiments.

**THEORY**

### Forward model

The forward modeling problem of finding traveltimes for curved rays in a layered Earth can be solved with ray shooting (Hardage, 1983). This technique consists of using Snell’s law to trace a ray downward from the source to the receiver. To trace a ray to a specific receiver, the initial angle of the ray is varied using the bisection method (Acton, 1990) until the ray passes sufficiently close to the receiver (within 1 cm). The general formula for the traveltime, \( t \), of a ray traveling along a path through a slowness field is

\[
t = \int s(\ell) \cdot d\ell,
\]

where \( \ell \) is the raypath and \( s \) is the slowness (defined as the reciprocal of velocity). For a layered system as shown in Figure 1, the traveltime for a ray with an initial angle of \( \theta_i \), hitting a receiver at depth \( z \) in layer \( j \), where all the layers have an identical thickness of \( h \) but different constant slownesses, \( s_n \), is

\[
t = \frac{s_1 h}{\cos \theta_i} + \sum_{n=2}^{j-1} \frac{s_n h}{\cos \left[ \arcsin \left( \frac{s_1}{s_n} \sin \theta_i \right) \right]} + \frac{\left[ z - (j-1)h \right]}{s_j} \left( 1 - \frac{s_1^2 \sin^2 \theta_i}{s_j^2} \right)^{2}
\]

\[
+ \left[ z - (j-1)h \right] s_1 \sin^2 \theta_i - s_j \left( 1 - \frac{s_1^2 \sin^2 \theta_i}{s_j^2} \right)^{2}.
\]

### Inversion

The traveltimes produced by equation (2) depend on slowness in a nonlinear way. Many different techniques can be used to invert nonlinear systems. We used a linearized Occam’s inversion scheme, which can be thought of as a first-order Taylor series expansion of the traveltime vector in slowness space (Menke, 1984). We used the “jumping” method (e.g., Scales et al., 1990) to solve the inverse problem. The jumping method allows us to apply constraints directly to the slowness model:

\[
s^* = (J_n^T J_n + \lambda^2 W_m W_m)^{-1} J_n^T (t^n + J_n s^{n-1}).
\]

In equation (3), \( s^* \) is the vector of layer slownesses at iteration \( n \), \( W_m \) is the model weighting matrix, a set of constraint equations that, in this case, minimize the first derivative of the layer slownesses in the vertical direction. The vector \( \Delta t^* \) is the desired change in the traveltimes, the difference between the forward-modeling results for iteration \( n \) and the data. The scalar \( \lambda \) controls the tradeoff between fitting the data and the model regularization. The Jacobian matrix \( J_n \) contains the derivatives of the forward-modeled traveltimes at iteration \( n \) with respect to the slownesses:

\[
J_{ij} = \frac{\partial t_i}{\partial s_j}
\]

These derivatives are obtained by differentiating equation (2). For the traveltime to receiver \( i \) at a depth of \( z \) in layer \( j \), these derivatives are

\[
\frac{\partial t_i}{\partial s_1} = \frac{h}{\cos \theta_i} + \sum_{n=2}^{j-1} \frac{h s_1 \sin^2 \theta_i}{s_n \left( 1 - \frac{s_1^2 \sin^2 \theta_i}{s_n^2} \right)^{2}}
\]

\[
+ \left[ z - (j-1)h \right] s_1 \sin^2 \theta_i - s_j \left( 1 - \frac{s_1^2 \sin^2 \theta_i}{s_j^2} \right)^{2}.
\]

\[
\frac{\partial t_i}{\partial s_n} = \frac{h}{\sqrt{1 - \left( \frac{s_1 \sin \theta_i}{s_j} \right)^2}} \frac{h s_n^2 \sin^2 \theta_i}{s_n^2 \left( 1 - \frac{s_1^2 \sin^2 \theta_i}{s_n^2} \right)^{2}}
\]

\[
\frac{\partial t_i}{\partial s_j} = \left[ z - (j-1)h \right] s_1 \sin^2 \theta_i - s_j \left( 1 - \frac{s_1^2 \sin^2 \theta_i}{s_j^2} \right)^{2}.
\]

(5)

Figure 1. Schematic diagram illustrating the parameters used for the curved-ray forward model. Each layer has a thickness of \( h \) and slowness of \( s_n \). The angle between the segment of the ray in layer \( n \) and the vertical is \( \theta_n \). The depth of the receiver is \( z \).
These analytical expressions can be evaluated using the slowness model of the previous iteration and the angles produced by the ray-shooting forward model.

At each iteration, the traveltimes to the receivers through the current slowness estimate are forward-modeled, $\mathbf{J}$ and $\Delta t$ are calculated, and the new slowness model is found using equation (3). The inversion is halted when the magnitude of the traveltimes residual does not decrease between iterations.

All linearized inversions require a starting model. Often, the choice of starting model can strongly influence the results (e.g., Vasco et al., 1996). Two methods for choosing a starting model are to try many starting models to characterize the non-uniqueness or to choose a starting model that prior information or preliminary analysis suggests is close to the “true” answer. While the first approach gives a more complete answer, the second is sufficient for our purposes. For the inversion described above, synthetic experiments have shown that starting models with velocities reasonably close to the true values produce accurate final-velocity models (Moret, 2003).

At many shallow sites, the P-wave velocity contrast across the water table is large (e.g., Miller and Xia, 1998). The depth of the water table is known because the depth to water in the borehole can be measured. In our vertical first-derivative regularization scheme, each row of $\mathbf{W}_m$ constrains the slowness contrast between two adjacent layers to be small. We can incorporate the available prior information about the water table into the inversion by changing $\mathbf{W}_m$ so that it does not constrain the velocity contrast across the layer boundary corresponding to the water table.

**Discrepancy principle**

Two parameters required by the inversion are an estimate of the variance of the traveltime noise and a value of $\lambda$. Morozov’s discrepancy principle (Morozov, 1984; Hansen, 1992) provides a rationale for choosing these parameters.

Wahba (1990) applied the discrepancy principle using the function $v(\lambda)$, defined as

$$v(\lambda) = \frac{\| \Delta t \|^2}{\text{trace}(\mathbf{I} - \mathbf{J}^T \mathbf{J} + \lambda^2 \mathbf{W}_m^T \mathbf{W}_m)^{-1} \mathbf{J}^T \mathbf{J})}, \quad (6)$$

where iteration $f$ is the final iteration of the inversion. The denominator of equation (6) is an estimate of the degrees of freedom of the data noise (Wahba, 1990). When the optimum $\lambda$ value, $\lambda^*$, is used, the value of $v(\lambda^*)$ is equal to the variance of the traveltime error, $\sigma_T$ (Wahba, 1990).

Hansen (1992) found that the behavior of $v(\lambda)$ could be used to estimate both $\lambda^*$ and $\sigma_T^2$. To obtain these estimates, we invert the data for many values of $\lambda$ and then plot the function $v(\lambda)$ versus $\lambda^{-1}$. At high $\lambda$ values, $v(\lambda)$ decreases steadily with decreasing $\lambda$. When $v(\lambda)$ reaches the noise level of the data, the velocity model cannot fit the data better without becoming much less smooth, so $v(\lambda)$ levels off at a plateau and does not decrease further until much smaller values of $\lambda$. The value of the function $v(\lambda)$ at the plateau should be equal to the variance of the noise in the data, $\sigma_T^2$ (Hansen, 1992). The optimum regularization parameter, $\lambda^*$, is the $\lambda$ value at the left end of the plateau, that is, the most highly regularized model with the desired value of $v(\lambda)$. Hansen (1992) suggested that higher $\lambda$ values could also be used to compensate for the inexactness of the Jacobian matrix.

**Uncertainty**

The final velocity model is more useful if the uncertainty in each layer’s velocity is known. The model covariance matrix contains information about the effect of data errors on the model parameters (Menke, 1984). The diagonal of the model covariance matrix gives the variances of each parameter. The off-diagonal terms are the covariances, which contain information on the relationships between parameters. As shown in Alumbaugh and Newman (2000), the model covariance matrix for a linearized inversion can be calculated using

$$C_m = \left[ \mathbf{J}^T \mathbf{J} + \lambda^2 \mathbf{W}_m^T \mathbf{W}_m \right]^{-1}. \quad (7)$$

In order to obtain the 95% confidence intervals for each layer’s slowness, we multiply the square root of each diagonal element of the matrix by 1.96 (95% of the area beneath a Gaussian curve is contained within 1.96 standard deviations of the mean). The uncertainties in the slownesses are normally distributed, so the uncertainties in the velocities are not. The 95% confidence limits for the layer velocities are

$$v_i^{\text{min}} = \frac{1}{s_i + 1.96\sigma_{i_{\text{model}}}}, \quad v_i^{\text{max}} = \frac{1}{s_i - 1.96\sigma_{i_{\text{model}}}}. \quad (8)$$

The equation for $v_i^{\text{max}}$ is discontinuous. If $\sigma_{i_{\text{model}}}$ is larger than $s_i$, then the lower 95% = 0 confidence limit for $s_i$ is negative, making $v_i^{\text{max}}$ negative also. This physically unrealistic negative value of $v_i^{\text{max}}$ indicates that no confidence can be placed in the estimate of $v_i$.

**SYNTHETIC EXAMPLE**

We tested our inversion routine on a synthetic data set. The test had three objectives: (1) to see whether the noise value estimated using the discrepancy principle was correct, (2) to verify that the inversion results were close to the original velocity model, and (3) to compare the results of our routine with a straight-ray inversion.

We generated the synthetic data using the ray-shooting algorithm described above. In the simulated survey, the shot was located 1 m away from the wells and the receivers were positioned every 0.1 m from just below the “water table” at 2 m depth to the bottom of the well, resulting in 179 synthetic traveltimes. We simulated picking errors in our data by adding Gaussian noise with a standard deviation of 0.05 ms.

We inverted this noisy data using many different $\lambda$ values to generate a discrepancy principle plot for these noisy synthetic traveltimes (Figure 2). The leftmost $\lambda$ on the plateau is 40000 (1/\(\lambda = 2.5 \times 10^{-4}\)). The plateau value corresponds to Gaussian error with a standard deviation of 0.045 ms, a 10% underestimate of the true value. This result suggests that the discrepancy principle can be used to estimate data error levels with reasonable accuracy.

Figure 3 shows the inversion result produced using a starting model with a velocity of 500 m/s above the water table and 2500 m/s below and a $\lambda$ value of 40000. The velocity across the synthetic water table was not constrained. The final velocity model has the same general structure and velocity values as the synthetic model but has less abrupt velocity contrasts because of model regularization.

We also tested the inversion using synthetic data with both Gaussian noise and a static time shift. The final velocity model
is nearly identical to Figure 3 except for a small change in the velocity of the layer above the water table (e.g., a change of <10 m/s for a shift of 0.1 ms). All of the rays travel nearly vertical paths in this layer, so a change in the layer velocity results in a nearly constant shift in the traveltimes. The discrepancy principle plot for the synthetic data with the static shift closely resembles Figure 2. This similarity indicates that noise estimates from discrepancy principle plots represent the random component of traveltimes.

We also used the synthetic experiment to compare our nonlinear curved-ray inversion to a linear straight-ray inversion. The straight-ray code used an inversion scheme similar to Lizarralde and Swift (1999):

$$s = (G^T G + \lambda^2 W_m^T W_m)^{-1} G^T t.$$  \hspace{0.5cm} (9)

The elements of the linear operator $G$ are the path lengths of each ray in each layer. Because of the straight-ray assumption, the path lengths do not depend on the layer slownesses, making the problem linear.

We chose a $\lambda$ value for the straight-ray inversion by using the $\chi^2$ statistic (Lizarralde and Swift 1999). The $\chi^2$ statistic compares the misfit of the solution with the expected misfit, given the traveltime noise. When the normalized $\chi^2$ statistic has a value near 1.0, the results fit the data with the expected degree of misfit. We systematically varied $\lambda$ and found that a $\lambda$ value of 10,000 produced a normalized $\chi^2$ value of 1.25. Because changing $\lambda$ slightly to get a normalized $\chi^2$ value closer to 1.0 would not have qualitatively changed the resulting velocity model, we used a $\lambda$ value of 10000 for the final inversion result.

Figure 3 compares the results for the straight-ray and the curved-ray inversions of the synthetic data. The straight-ray velocity model is more oscillatory. Because of the simplifying straight-ray assumption, the straight-ray inversion results must include unrealistic velocity changes to reach the desired traveltime misfit level. The larger scale variations in the straight-ray inversion result match the synthetic model at deeper layers, where the straight-ray assumption is a better approximation. However, the straight-ray inversion is unable to match the layer velocities above roughly 7 m. At shallow depths, the raypaths cannot be reasonably approximated by straight rays, so the inversion cannot recover the true velocity model. While ignoring the $\chi^2$ information and choosing a larger $\lambda$ value would eliminate the oscillations in the straight-ray inversion results, a larger $\lambda$ value would not be able to overcome the fundamental limitations of the forward model in dealing with large velocity contrasts within a few meters of the surface. The results of this comparison suggest that curved-ray inversion is needed to accurately recover the subsurface velocity structure from traveltimes for typical near-surface survey geometries and velocities.

FIELD EXAMPLE

We used the curved-ray inversion routine described above to invert a shallow reverse vertical seismic profile (RVSP). An RVSP uses a surface geophone to record energy from a borehole source. The principle of reciprocity holds that the traveltime for a ray remains unchanged if the source and receiver locations are reversed, so the same inverse code can be used for VSPs and RVSPs.

The RVSP data set used in this study was collected in well C4 at the Boise Hydrogeophysical Research Site (BHRS), a research wellfield near Boise, Idaho (Barrash and Knoll, 1998). The BHRS is developed in a shallow, unconsolidated, sand-and-cobble aquifer underlain by a clay aquitard. Because preferential flow paths play an important role in flow and transport processes in sand-and-cobble deposits (e.g., Jussel et al., 1994), considerable interest in mapping the heterogeneity of such sediments exists.

A borehole sparker was used as the seismic source for the RVSP. A sparker is a rubber-walled cell full of salt water containing a spark gap. When a high voltage (approximately 4 kV) is discharged across the spark gap, the ionization of a small amount of salt water caused by the electric current creates an
acoustic disturbance that expands to the rubber walls. The rubber walls transmit the disturbance to the surrounding fluid in the form of P-wave energy. In dry holes, the coupling between the sparker walls and the formation is poor, so data collected using the sparker has the highest signal-to-noise ratio below the water table. We used a vertical-component geophone to record the arriving seismic energy. The resonance frequency of the geophone was 10 Hz and, according to the manufacturer, it had a uniform amplitude response to roughly 1 kHz. The geophone was placed less than a meter from the well, and the elevation of the geophone was surveyed. The sparker was lowered into the well and fired every 0.1 m from just below the water table to the bottom of the well.

The data were recorded on a 24-bit engineering seismograph with a 0.0625 ms sample interval and a 64 ms record length. A 24 dB acquisition gain and an antialias filter were applied. The traces collected are shown in Figure 4. The first arrival traveltimes of the data were picked with a commercial seismic processing package that uses subsample interpolation. We used a total of 121 traveltimes.

The data were inverted using the routine described above. At the time of data collection, the water level in the well was 1.62 m below land surface. The large velocity contrast could be slightly higher than the water level in the well due to capillary effects. At the BHRS, however, the height of the capillary fringe would be relatively small because of the coarse grain size of the sediments in the aquifer (Dingman, 1994). We used borehole deviation logs to set a different horizontal offset for each shot. The model weighting matrix $W_m$ was designed to prevent it from constraining the velocity contrast across the water table. The starting velocity model we used was 2500 m/s below the water table and 300 m/s above the water table. This model was based on the results of a seismic refraction experiment at the BHRS (S. Goldstein, 2002, personal communication).

The discrepancy principle plot for the inversion of the C4 data set is shown in Figure 5. The slope of the curve decreases with decreasing $\lambda$, but the curve does not have a distinct plateau. This continuous decrease may indicate that the noise level is not uniform. Because the signal amplitude decreases with depth, the traveltine noise may increase with depth. We chose to use the highest $\nu(\lambda)$ value in the region of slowly decreasing $\nu(\lambda)$ to avoid underestimating the uncertainties in the layer velocities. This value corresponds to a $\lambda$ value of 25000 and a $\sigma_T$ of 0.027 ms. The choice of $\lambda$ is somewhat arbitrary, but using nearby $\lambda$ values does not noticeably affect the results. The $\sigma_T$ value is between one-half and one-third of the sample interval, an error level similar to other studies that have used subsample picking (e.g., Lanz et al., 1998).

Figure 6 shows the inversion results and 95% confidence intervals. The confidence intervals are large above 4 m because of the lack of data and below 16 m because of the low ray coverage. As a result, little confidence can be placed in the velocity model except between 4 and 16 m.

To verify our results, we compared the final slowness model produced by the inversion with a neutron porosity log from well C4 (Figure 7). Details concerning the collection and processing of the neutron log can be found in Barrash and Clemo (2002). While seismic slowness generally increases as porosity increases, we would not necessarily expect a simple relationship between the two, as other factors in addition to porosity (e.g., grain-size distribution, pore-space geometry, effective stress) can influence the seismic properties of a medium.
First-order similarity is evident between the seismic slowness and porosity profiles (Figure 7) with zones of relatively high values centered at about 5, 13, and 14.5 m. Examination of core from wells at the BHRS indicates that such higher porosity zones occur where the proportion of sand matrix is higher relative to the cobble framework clasts and/or where packing of the sediment is less dense (Reboulet, 2003). Here we note that the slowness peak at 7 m does not correspond to a high-porosity zone (Figure 7), and core analysis indicates sediments with low proportion of sand matrix in this interval. In general, however, changes in porosity are reflected in the slowness model. This suggests that our inverse routine was largely successful in imaging variations in subsurface properties.

CONCLUSIONS
Ray bending affects shallow VSPs and RVSPs more than deeper VSPs because of the geometry of acquisition and the strong velocity contrasts in the shallow subsurface. This ray bending must be considered when inverting shallow VSP or RVSP traveltimes to produce accurate velocity models. In the synthetic example, the straight-ray inversion was unable to reproduce the true velocity model above 7-m depth. The velocity model produced by the straight-ray inversion also required high-frequency oscillations to satisfy the data. In practice, anyone using a straight-ray inversion for such a data set would use a higher $\lambda$ value, either ignoring $\chi^2$ or increasing their estimate of $\sigma_d$. Using a higher $\lambda$ value would reduce the resolution of the inversion.

The synthetic experiment also showed that the curved-ray inversion presented here can produce accurate velocity models in the presence of high velocity contrasts. Of particular interest was the use of the discrepancy principle to estimate $\sigma_d$ and to choose $\lambda$. The $\sigma_d$ estimate was within 10% of the true noise level and the chosen $\lambda$ value resulted in an accurate velocity model.

The plateau in the plot of $\nu(\lambda)$ was not as distinct for the field example. We suspect that the more gradual decline of $\nu(\lambda)$ was due to variation in the noise level with depth. Because we did not wish to underestimate the uncertainty in the final velocity model, we chose the highest possible “plateau” value. The final velocity model for a shallow RVSP data set was consistent with prior information. The uncertainties associated with the velocities show that the inversion of VSP or RVSP traveltimes can distinguish between zones of different velocity on a scale of ~1 m. The ability to image the subsurface at this level of resolution may be very helpful in engineering and hydrologic investigations.

ACKNOWLEDGMENTS
We would like to thank Lee Liberty and Paul Michaels for their input into this project. Steve Shedlock of SubTerra Surveys provided the sparker sonde used in this experiment. We are also grateful to the associate editor and several reviewers for their comments. The Army Research Office (DAAD19-00-1-0454 and DAAH-96-1-0318) and the Environmental Protection Agency (X-97008501-0) funded this work.

REFERENCES